JIPB Journal of Integrative Plant Biology

Evolution and protein interactions of AP2 proteins in Brassicaceae: Evidence linking development and environmental responses

Liping Zeng[1†], Yue Yin[1†], Chenjiang You[1], Qianli Pan[1], Duo Xu[1], Taijie Jin[1], Bailong Zhang[1*] and Hong Ma[1,2*]

[1]Ministry of Education Key Laboratory of Biodiversity and Ecological Engineering, Institute of Plant Biology, Institute of Biodiversity Sciences, Center for Evolutionary Biology, School of Life Sciences, Fudan University, Shanghai 200433, China, [2]Institutes of Biomedical Sciences, Fudan University, Shanghai 200433, China. [†]These authors contributed equally to this work. *Correspondences: Hongma@fudan.edu.cn (Ma is fully responsible for distributions of all materials associated with this article); bailong.zhang@ucr.edu

Abstract Plants have evolved a large number of transcription factors (TF), which are enriched among duplicate genes, highlighting their roles in complex regulatory networks. The *APETALA2/EREBP-like* genes constitute a large plant TF family and participate in development and stress responses. To probe the conservation and divergence of *AP2/EREBP* genes, we analyzed the duplication patterns of this family in Brassicaceae and identified interacting proteins of representative *Arabidopsis* AP2/EREBP proteins. We found that many *AP2/EREBP* duplicates generated early in Brassicaceae history were quickly lost, but many others were retained in all tested Brassicaceae species, suggesting early functional divergence followed by persistent conservation. In addition, the sequences of the AP2 domain and exon numbers were highly conserved in rosids. Furthermore, we used 16 *A. thaliana AP2/EREBP* proteins as baits in yeast screens and identified 1,970 potential AP2/EREBP-interacting proteins, with a small subset of interactions verified *in planta*. Many *AP2* genes also exhibit reduced expression in an anther-defective mutant, providing a possible link to developmental regulation. The putative AP2-interacting proteins participate in many functions in development and stress responses, including photomorphogenesis, flower development, pathogenesis, drought and cold responses, abscisic acid and auxin signaling. Our results present the *AP2/EREBP* evolution patterns in Brassicaceae, and support a proposed interaction network of *AP2/EREBP* proteins and their putative interacting proteins for further study.

Keywords: APETALA2; EREBP; evolution; Brassicaceae; protein interaction
Citation: Zeng L, Yin Y, You C, Pan Q, Xu D, Jin T, Zhang B, Ma H (2015) Evolution and protein interactions of AP2 proteins in Brassicaceae: Evidence linking development and environmental responses. **J Integr Plant Biol** XX:XX–XX doi: 10.1111/jipb.12439
Edited by: Chung-Mo Park, Seoul National University, Korea
Received Sept. 11, 2015; **Accepted** Oct. 15, 2015
Available online on Oct. 16, 2015 at www.wileyonlinelibrary.com/journal/jipb

INTRODUCTION

Plants live in ecosystems with ever changing environments; accordingly, plants have evolved many specific features that facilitate the interactions with their environments. Thus plant-specific functions control complex cellular networks that enable their adaptation to various environmental conditions (Arabidopsis Interactome Mapping 2011). Some of the plant-specific functions are carried out by gene families encoding transcription factors, which are key components of comprehensive regulatory and interactive networks mediating plant-environment interactions (Mizoi et al. 2012). During evolution, plant genes have experienced gene and domain duplication and the duplicate genes could then differentiate in terms of expression patterns and interaction partners (Kim et al. 2006). Thus, the study of the phylogeny and interaction patterns of the plant-specific transcription factor gene family could provide important clues about plant development and growth. In particularly, many groups of plants have experienced whole-genome duplication(s) (WGDs) (e.g., Jiao et al. 2011), but the patterns of gene duplication and loss in genes for transcriptional factors following a WGD have not been examined in detail.

The *APETALA2-like (AP2-like)* and *Ethylene-Responsive Element-binding Proteins/Factors (EREBP/ERF)* genes form a large plant-specific family encoding proteins that contains one or two DNA-binding domain(s) called the AP2 domain (Weigel 1995; Okamuro et al. 1997; Riechmann and Meyerowitz 1998). *AP2/EREBP* genes comprise *AP2-like* genes (with two AP2 domains) and *EREBP-like* or *ERF-like* genes (with one AP2 domain) (Riechmann and Meyerowitz 1998). The *AP2-like* genes form the *AP2* subfamily, which can be further divided into the *ANT* group and the *AP2* group in seed plants (Shigyo and Ito 2004; Mizoi et al. 2012). The *EREBP-like* genes include *DREB*, *ERF* and *RAV* (with another DNA-binding domain called B3) subfamilies. Moreover, the *DREB* and *ERF* subfamilies each have six subgroups, which are named, respectively, *DREB-A1* to *DREB-A6* and *ERF-B1* to *ERF-B6* (Sakuma et al. 2002; Mizoi et al. 2012).

AP2/EREBP genes participate in many aspects of plant development and stress responses (Kim et al. 2006; Mizoi et al. 2012). Most of the known functions of *AP2-like* genes are important for developmental processes, whereas the *EREBP-*

like genes are primarily known for regulating stress responses (Riechmann and Meyerowitz 1998; Kim et al. 2006). For example, the *Arabidopsis thaliana AP2* gene has many important functions in reproductive development, including the specification of floral organ identities and control of floral homeotic gene expression (Bowman et al. 1989; Kunst et al. 1989), regulation of flowering time and floral meristem (Bowman et al. 1991; Coen and Meyerowitz 1991), and modulation of seed development (Jofuku et al. 1994; Jofuku et al. 2005; Ohto et al. 2005). Also, homologs of *AP2* called *TARGETS OF EAT* (*TOE*) affect flowering time in *A. thaliana* by repressing the expression of flowering regulatory genes *FLOWER LOCUS T* (*FT*) and *Suppressor of Overexpression of Constans 1* (*SOC1*) (Krizek 2003; Mathieu et al. 2009; Yant et al. 2010). The *TOE* genes include *TOE1, TOE2, TOE3, SCHLAFMÜTZE* (*SMZ*) *and SCHNARCHZAPFEN* (*SNZ*) (Aukerman and Sakai 2003) and form a small subgroup with *AP2*. These *AP2* subgroup genes are negatively regulated by *microRNA172* (*miR172*) at the post-transcriptionally (Aukerman and Sakai 2003; Schmid et al. 2003; Chen 2004). We recently showed that the TOE proteins could physically and genetically interact with the key flowering regulator CONSTANS (CO) and repress the transactivation ability of CO, thereby preventing premature flowering (Zhang et al. 2015). In addition, the *ANT* gene, the namesake for the *ANT* group of the *AP2-like* subfamily, regulates ovule development and floral organ growth (Elliott et al. 1996; Klucher et al. 1996).

Members of the *EREBP* subfamily have been implicated in response to environmental stresses (Kizis et al. 2001; Mizoi et al. 2012). Specifically, *ERF* genes are regulated by hormones and environmental factors such as cold, drought or high salinity (Fujimoto et al. 2000). *ERF1, ERF2* and *ERF5* can bind to the GCC-box and act as transcriptional activators, whereas *ERF3* and *ERF4* act as repressors in *A. thaliana* (Fujimoto et al. 2000). Also, the *C-Repeat Binding Factor* (*CBF*) genes encoding DREB type proteins act as transcriptional activators of genes important for freezing tolerance in response to cold. The tandem duplicates *DREB1A/CBF3, DREB1B/CBF1* and *DREB1C/CBF2* are also induced by cold (Stockinger et al. 1997; Gilmour et al. 1998; Medina et al. 1999). Some *EREBP* members are developmental regulators: *A. thaliana RAV1* is downregulated by brassinosteroid and may act as a negative regulator during plant development (Hu et al. 2004). Further, *DREB-A6* (also named *WOUND INDUCED DEDIFFERENTIATION 1* (*WIND1*)) is involved in cell dedifferentiation in *A. thaliana* (Iwase et al. 2011).

The phylogeny of the *AP2/EREBP* family and the pattern of changes in gene structure and protein domain among flowering plants are crucial for elucidating the evolution of this developmentally and physiologically important gene family. The AP2/EREBP domain has not been found outside the plants, so it was considered as plant-specific (Riechmann and Meyerowitz 1998; Krizek 2003), or its homologs in non-plant organisms have diverged too much to be detectable by sequence comparison. Previous phylogenetic studies classified the *AP2/EREBP* family into subfamilies and proposed a model for the evolution of domain organization in the *AP2* subfamily (Kim et al. 2006), whose members have two AP2 domains. However, the relationships of *AP2/EREBP* genes among members of a plant family and the patterns of gene retention following a WGD remain unclear. Specifically, it is not clear whether the same duplicate(s) generated in an ancestor tend to be lost in all/many species or only lost in some species but retained in others.

Brassicaceae is a large family of eudicot angiosperms with many economically important plants, such as cabbage and oilseed rape, as well as the model plant *A. thaliana*. There is strong evidence for two WGDs at or near the origin of this family (Blanc et al. 2003); in addition, the genome sequences of several members of this family have been determined (Dassanayake et al. 2011; Hu et al. 2011; Wang et al. 2011; Slotte et al. 2013), providing an excellent opportunity to examine the gene duplication and loss patterns of the *AP2/EREBP* family. To investigate the *AP2/EREBP* phylogeny in Brassicaceae, we searched extensively for *AP2/EREBP* homologs in Brassicaceae and performed phylogenetic analyses of each subfamily. We investigated retention patterns of duplicated *AP2/EREBP* genes during Brassicaceae evolution and the conservation and variation of AP2 domain sequences and exon numbers among subfamilies. Furthermore, as a test for functional similarity and differences among different AP2/EREBP proteins, we searched for potential interaction partners of selected representative *Arabidopsis* AP2/EREBP proteins, using yeast two-hybrid screening. We detected 1,970 candidate interaction proteins of 16 AP2/EREBP proteins, of which 300 highly interconnected proteins form the core AP2/EREBP interactome. Finally, as an example of possible regulatory interaction with other regulatory genes, we examined the expression of *AP2* family members in terms of possible effect of anther developmental regulators.

RESULTS

Identification of AP2/EREBP genes

To investigate the evolution of the *AP2/EREBP* family in Brassicaceae, we identified *AP2/EREBP* genes by performing HMMER to search for the predicted proteins from five representative species of Brassicaceae and two other rosids species (as outgroups), whose genomes have been completely sequenced. After further confirmation using BLAST and SMART searches, we finally identified 144, 143, 144, 281, 140, 109 and 209 *AP2/EREBP* genes from *A. thaliana*, *A. lyrata*, *Capsella rubella*, *Brassica rapa*, *Eutrema salsugineum*, *Carica papaya* (Caricaceae) and *Populus trichocarpa* (Salicaceae), respectively. The numbers of *AP2/EREBP* genes in *A. thaliana* (144) and *P. trichocarpa* (209) were very close to previously reported findings (147 and 200). Except *B. rapa*, which experienced a lineage-specific recent WGD event and has large numbers of *AP2/EREBP* genes, other Brassicaceae species contained similar numbers of *AP2/EREBP* genes.

Duplicated copies of *AP2/EREBP* family were partially or completely lost in the MRCA of Brassicaceae

Besides the conserved AP2 domain (containing about 50–60 amino acids, hereafter aa), the sequences of AP2/EREBP proteins are highly variable. We initially constructed a preliminary phylogenetic tree (ML tree, using a maximum likelihood method) based on alignment of full-length sequences of 1,170 AP2/EREBP proteins from the seven species mentioned above and then used the *A. thaliana* subgroup definition (Dietz et al. 2010), to divide the sequences

in other six species into 14 subfamilies: *AP2*, *RAV*, *ERF-B1* to *ERF-B6*, and *DREB-A1* to *DREB-A6* (Kim et al. 2006). Because of the extreme variation in regions flanking the AP2 domain among subfamilies, the initial alignment with 1,170 sequences were divided into 14 clusters according to subfamily classification for generating more accurate alignments.

To investigate the evolutionary history of subfamilies in *AP2/EREBP* family, ML trees of subfamilies were reconstructed. These subfamilies experienced uneven expansion according to the numbers of genes in the most recent common ancestor (MRCA) of rosids (Figures 1, S1–S12). We found that most expansions of *AP2/EREBP* family genes had occurred before the divergence of Brassicaceae and WGD likely played an important role in the expansion of *AP2/EREBP* family in Brassicaceae (Figures 1, S1–S12). PGDD analyses found that 83 genes in *A. thaliana* were duplicates probably due to WGD. Moreover, likely tandem duplication has been identified in the MRCA of Brassicaceae or specific lineages in several subfamilies. In the *DREB-A1* subfamily, lineage-specific tandem duplication was associated with *CBF* homologs *AT4G25470*, *AT4G25480* and *AT4G25490* (Stockinger et al. 1997; Medina et al. 1999) in *A. thaliana*, *C. rubella* and *B. rapa* (Figure 1). In the *DREB-A2* subfamily, the pair of AT2G40340 and AT2G40350 resulted from an *A. thaliana*-specific tandem duplication (Figure S1). In addition, prior to the divergence of the Brassicaceae species sampled here, tandem duplication events in *ERF-B3* subfamily resulted in multiple clusters. In *A. thaliana* these include (i) AT3G23220, AT3G23230 and AT3G23240, (ii) AT4G17490 and AT4G17500, (iii) AT5G47220 and AT5G47230, and (iv) AT5G61590 and AT5G61600. Their orthologs are also retained in four other species of Brassicaceae examined here (Figure S7). On the other hand, two *A. thaliana* tandem genes, AT5G67190 and AT5G67180, have highly different sequences (Figure S3). Whereas AT5G67190 contains only one AP2 domain and belongs to the *DREB-A5* subfamily, AT5G67180 (also called *TOE3*) has two AP2 domains and belongs to *AP2* subfamily (Figures S3, S12), suggesting they are not the result of recent duplication.

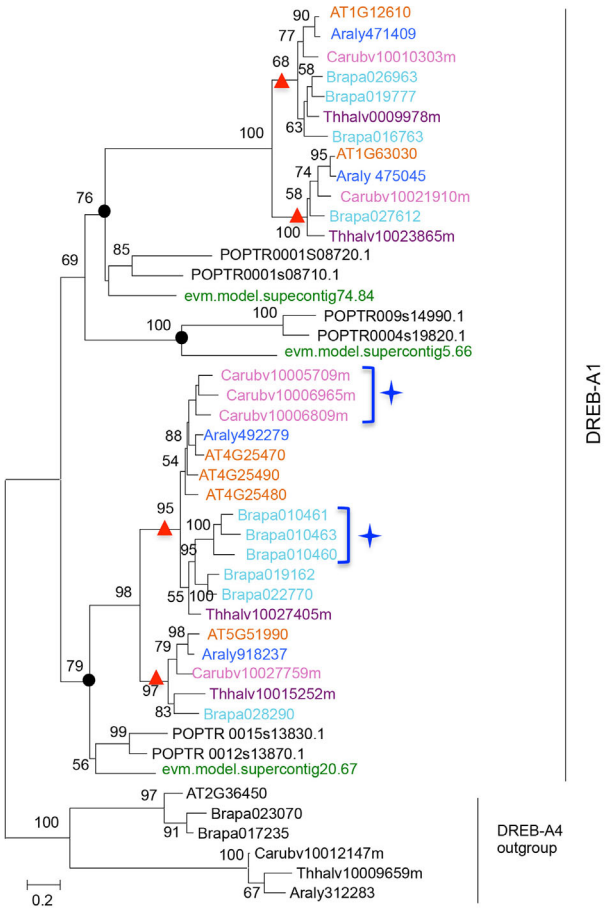

Figure 1. A phylogenetic tree of the *DREB-A1* subfamily with several *DREB-A4* genes as outgroup

Bootstrap supports (>50) obtained by RAxML are shown. Gene IDs of each species are shown in different colors. Black circles represent that the common ancestor of rosids contain three ancestor *DREB-A1* genes; blue stars stand for tandem duplications in *Arabidopsis thaliana*, *Capsella rubella* and *Brassica rapa*, respectively; red triangles point to retained genes after two rounds of WGDs in early Brassicaceae.

The MRCA of Brassicaceae has experienced two rounds of WGD (which are called alpha and beta) (Blanc et al. 2003), such that members of the *AP2/EREBP* family should form clades of four close copies if none of the duplicates due to the WGDs were lost. However, no such four-copy clades were detected in any subfamilies or any of the Brassicaceae species examined here, suggesting that many of the duplicate *AP2/EREBP* copies were probably lost prior to the divergence of the Brassicaceae species examined here, early in the history of this family.

The AP2 domains in EREBP subfamilies show high conservation

It was reported that the two AP2 domains found in the *AP2* clade resulted from a duplication prior to the divergence of euAP2 and ANT groups (Kim et al. 2006). We focus on the conservation and diversification the *EREBP* genes with only one AP2 domain. Unlike the highly variable sequences of the regions outside the AP2 domain of orthologs among *A. thaliana*, *C. papaya* and *P. trichocarpa*, sequences of AP2 domain in subfamilies were strikingly similar and several residues were highly conserved. To examine the conservation and variation of AP2 domain in the *DREB* and *ERF* subfamilies, we aligned the AP2 domain of proteins from the five Brassiciceae species, *C. papaya* and *P. trichocarpa* and generated amino acids logos for each *DREB* and *ERF* subfamily (Figure 2). Except the *ERF-B3* subfamily, which has an insertion of one amino acid, the length of AP2 domain in all remaining *DREB* and *ERF* subfamilies is 52 aa (Figure 2). The AP2 DNA binding domain is conserved in Brassicaceae history: 12 residues are entirely identical (100%) and 13 residues are highly conserved (80%–90%) among all *DREB* and *ERF* subfamilies.

The high degree of amino acid sequence similarity of AP2 domain of DREB and ERF groups in Brassicaceae suggests that the EREBP proteins might be conserved in DNA binding; in contrast, the highly variable sequences of the regions flanking the AP2 domain, implies that members of different subfamilies might differ in their interacting proteins. Strikingly, the ERF-B3 group has an insertion of one amino acid, with potentially different DNA binding properties from other EREBPs.

Figure 2. Amino acid sequences of the AP2 domain in *DREB* and *ERF* subfamilies
Blue rectangle marks the 1-aa insertion of the *ERF-B3* subfamily. For all other subfamilies, the length of the AP2 domain is 52 aa. Red stars indicate twelve completely identical (100%) residues and green needles indicate thirteen highly conserved (80%–90%) residues.

AP2-like genes possess multiple exons but EREBPs have single or few exons

Next we examined the exon numbers and found that exon numbers displayed similar patterns between members within each subfamily in Brassicaceae and even in the rosid species we examined, but were dramatically different between one-AP2-domain lineages and the two-AP2-domain lineage. In the *AP2* subfamily, whose members contain two AP2 domains, all genes contained more than five exons in the five Brassicaceae species and *P. trichocarpa*. However, most genes in the *RAV*, *DREB* and *ERF* subfamilies, whose members have a single AP2 domain (Kim et al. 2006; Mizoi et al. 2012), have only one exon (Figures 3, S13). All *RAV* genes lack intron in Brassicaceae species, and only two *RAV* genes have introns in *P. trichocarpa*. Besides one *A. lyrata* and one *C. papaya* gene, nearly all *ERF-B2* genes have two exons in Brassicaceae species and *P. trichocarpa*. Taken together, these patterns suggested that single or limited exon was the common feature of one-AP2-domain subfamilies, while genes with multiple exons were the characteristic of the *AP2* subfamily (two AP2 domains) in the MRCA of Brassicaceae. The results here suggest that the ancestral genes for the one-AP2-domain subfamilies probably had none or one intron; one possibility is that the common ancestor of all *AP2* genes had more introns, but all or most introns were lost after the one-AP2-domain clade separated from the two-domain clade. Because retrotransposition is thought to be a major mechanism for gene duplication resulting in intronless genes, it is possible that some of the single AP2-domain genes were duplicated via retrotransposition.

Our results on gene structure support the previous classification of the AP2 and EREBP groups using the number of the AP2 domain (Okamuro et al. 1997). The two-AP2-domain genes have multiple exons, whereas the one-AP2-domain genes have one or two exons, with only few exceptions with multiple exons.

Screening for interactive proteins of representative AP2/EREBP proteins

Although a large-scale protein-protein interaction map of the *A. thaliana* has uncovered a large number of potential interactive partners, which linked novel functional pathways (Arabidopsis Interactome Mapping 2011), the *AP2* family was not a specific focus of the study and few interactive proteins were identified for AP2 proteins. To investigate possible regulatory mechanisms used by AP2/EREBP proteins using interactive proteins, we selected 56 representative AP2/EREBP proteins as baits for yeast two-hybrid (Y2H) screening to identify potential interacting proteins. These 56 proteins were selected using phylogenetic and functional criteria: (i) one to two members were selected from each clade, and (ii) those with known functions were favored. For example, AP2 and TOE1 were from the AP2 subgroup, and ANT and At1g79700 were from the ANT subgroup. Among these 56 selected AP2-domain proteins, 26 showed strong auto-activation in yeast, two showed severe toxicity to yeast cells, and seven were not tested for autoactivation due to technical reasons; these were not used in subsequent analyses (Table 1; Figure S14). The remaining 21 proteins were used to screen for interactive proteins (Table 1).

Next, we performed Y2H screening of the 21 non-autoactivation bait proteins. Normalized *Arabidopsis* cDNA libraries generated from mRNA of various tissues and developmental stages were screened and the screens resulted in positive clones for 16 of the bait proteins, including TOE1 (AT2G28550) (Aukerman and Sakai 2003;

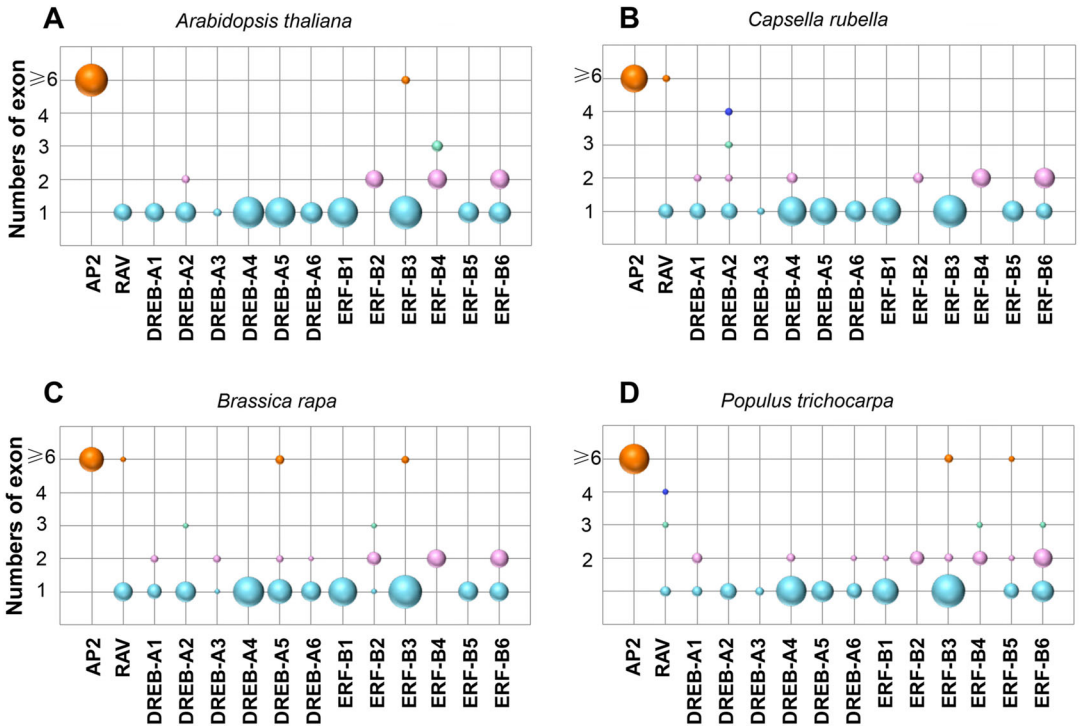

Figure 3. Exon numbers of each *AP2/EREBP* subfamily genes in (A) *Arabidopsis thaliana*, (B) *Capsella rubella*, (C) *Brassica rapa* and (D) *Populus trichocarpa*

The circles represent genes with the number of exons in different subfamilies. The exon numbers are shown on the Y-axis, with circles of different colors. The size of circles represents the percentage of gene numbers that contains different numbers of exon in each *AP2/EREBP* subfamilies.

Zhang et al. 2015), AP2 (AT4G36920, weakly auto-activating but chosen because of its importance as a developmental regulator) (Jofuku et al. 1994; Jofuku et al. 2005; Ohto et al. 2005; Yant et al. 2010), WRI4 (AT1G79700) (Iwase et al. 2011) and DREB-A2 protein AT5G18450, DREB-A4 proteins AT1G71450 and AT2G44940, DREB-A5 protein AT5G21960, DREB-A6 protein AT1G64380, ERF-B1 proteins ERF4 (AT3G15210) (Yang et al. 2005), ERF8 (AT1G53170) and ERF12 (AT1G28360), ERF-B3 proteins ERF2 (AT5G47220), AT5G51190 and AT2G20880 and ERF-B6 protein ESE3 (AT5G25190) (Mizoi et al. 2012) and SHN2 (AT5G11190) (Table 2). The numbers of identified prey proteins showed different patterns: the two-AP2-domains proteins had more preys, whereas the one-AP2-domain proteins had many fewer preys (Table 2). The total number of preys for the 16 baits was 1970, including 252 transcription factors, 215 proteins with functions for protein modification, 335 with protein processes, 20 for epigenetics related processes, 853 with other functions and 295 unknown function proteins (Figure 4A and Table 2). Taken together, our results have identified putative interaction proteins for representative AP2/EREBP family proteins.

TOE1 and related proteins showed similar interaction patterns

As a first step to eliminate false positives, the interactions identified by Y2H were retested in yeast cells. We cloned some of the full-length prey cDNAs and transformed the plasmids containing the prey cDNAs into yeast to validate the interactions of baits and preys (Figure S15A). In addition, as a case study for specificity of interactions, we selected 28 TOE1 interacting proteins, and tested their interactions with TOE1 and three other closely related proteins: TOE2, TOE3 and SNZ (Figures 5, S15B). We found that most of the TOE1 interacting proteins could also interact with TOE2, TOE3 and SNZ, with similar patterns of interaction strengths to that with TOE1. However, SNZ was slightly different from the other three bait proteins in that SNZ could not interact with AALP and EMB1374. In particular, COL1 had a strong interaction with TOE1, TOE2 and TOE3 but only a very weak interaction with SNZ. Also, all of the proteins we identified here that are involved in protein modification could interact with the four TOE-like proteins. Strikingly, the F-box protein ZTL, which was reported to control the circadian clock and influence flowering time (Somers et al. 2000; Somers et al. 2004; Takase et al. 2011), could interact with TOE-like proteins in our experiments (Figure 5). Based on the genetic experiments showing functions of *TOE-like* genes in regulation of flowering time, we surmised that these interactions might provide new insights into flowering time regulation. In conclusion, these results showed that several TOE1 paralogs have similar interaction patterns, suggesting that they are functionally conserved.

Test for *in vivo* interactions by BiFC

To examine whether the interactions identified by Y2H reflect *in planta* interactions, we used Bimolecular

Table 1. AP2/EREBP genes selection and autoactivation test

Gene ID	Name	Subfamily	Auto-activation	Gene ID	Name	Subfamily	Auto-activation
AT2G28550	*TOE1*	*AP2*	–	*AT1G50640*	*ERF3*	*ERF-B1*	+
AT4G36920	*AP2*	*AP2*	+/–	*AT1G53170*	*ERF8*	*ERF-B1*	–
AT1G79700	*WRI4*	*AP2*	–	*AT5G13910*	*LEP*	*ERF-B1*	+/–
AT1G13260	*RAV1*	*RAV*	–	*AT5G44210*	*ERF9*	*ERF-B1*	+/–
AT4G37750	*ANT*	*ANT*	+/–	*AT1G72360*	*ERF73*	*ERF-B2*	Nd
AT1G72570		*ANT*	Nd	*AT2G47520*	*ERF71*	*ERF-B2*	Nd
AT2G41710		*AP2*	Nd	*AT3G14230*	*RAP2.2*	*ERF-B2*	Tx
AT4G25480	*DREB1A*	*DREB-A1*	+	*AT3G16770*	*EBP*	*ERF-B2*	+
AT1G63030	*DDF2*	*DREB-A1*	+	*AT4G34410*	*RRTF1*	*ERF-B3*	–
AT4G25490	*CBF1*	*DREB-A1*	+	*AT2G44840*	*ERF13*	*ERF-B3*	+
AT5G18450		*DREB-A2*	Nd	*AT3G23220*	*ESE1*	*ERF-B3*	Nd
AT2G40340	*DREB2C*	*DREB-A2*	Tx	*AT3G23240*	*ERF1*	*ERF-B3*	+
AT5G18450		*DREB-A2*	–	*AT5G07580*		*ERF-B3*	Nd
AT2G40220	*ABI4*	*DREB-A3*	+	*AT5G47220*	*ERF2*	*ERF-B3*	–
AT5G25810	*TINY*	*DREB-A4*	+	*AT5G51190*		*ERF-B3*	–
AT5G52020		*DREB-A4*	–	*AT5G13330*	*RAP2.6L*	*ERF-B4*	+/–
AT3G16280		*DREB-A4*	+/–	*AT5G50080*	*ERF110*	*ERF-B4*	+
AT1G71450		*DREB-A4*	–	*AT4G23750*	*CRF2*	*ERF-B5*	+
AT1G33760		*DREB-A4*	+/–	*AT5G53290*	*CRF3*	*ERF-B5*	+
AT2G44940		*DREB-A4*	–	*AT1G25470*	*CRF12*	*ERF-B6*	+
AT1G71520		*DREB-A5*	–	*AT1G25470*	*CRF12*	*ERF-B6*	+
AT1G74930	*ORA47*	*DREB-A5*	+	*AT5G11190*	*SHN2*	*ERF-B6*	–
AT4G36900	*DEAR4*	*DREB-A5*	–	*AT5G19790*	*RAP2.11*	*ERF-B6*	+
AT5G21960		*DREB-A5*	–	*AT5G25190*	*ESE3*	*ERF-B6*	–
AT1G64380		*DREB-A6*	–	*AT1G49120*	*CRF9*	*ERF-B6*	+
AT1G78080	*WIND1*	*DREB-A6*	+	*AT5G67000*		*ERF-B6*	–
At3g15120	*ERF4*	*ERF-B1*	–	*At2g20880*	ERF53		–
AT1G28360	*ERF12*	*ERF-B1*	–	*AT3G61630*	*CRF6*		–

+, strong auto-activation; –, no auto-activation; +/–, weak auto-activation; Nd, not detect; Tx, toxicity in yeast cells.

Fluorescence Coupling (BiFC) in tobacco cells to test for interaction of 16 selected pairs of interactive proteins identified using yeast. We found most of the yeast interactions were reproduced in plant cells including the interactions of WRI4 with AT1G01060, STO (Yan et al. 2011), HAP5C, COL3 (Datta et al. 2006), and AT2G43770. Also reproduced are interactions of AFO and AT1G77080; AT5G63970 with AT1G71450; AT1G06760 with AT1G64380; AT1G18340 with AT1G64380; AT1G29150 with AT1G28360; AT2G20400 with AT1G28360, and that of AT3G51880 with AT1G71450 (Figure 6). Strikingly, all YFP signals for these interactions were observed in nuclei, consistent with AP2 members being known or putative transcription factors. However, three pairs of interactions failed to occur in plant cells, including those between AT1G80420 and AT1G64380, between AT5G01920 and AT1G28360, and between AT5G64920 and AT1G64380 (Figure 6). These results indicated that most of the interactions between partners identified in yeast cells could also have occurred in plant cells.

Table 2. AP2/EREBP proteins used for Y2H screening

Gene ID	Name	Subfamily	Prey No.
AT2G28550	*TOE1*	*AP2*	383
AT4G36920	*AP2*	*AP2*	421
AT1G79700	*WRI4*	*AP2*	341
AT5G47220	*ERF2*	*ERF-B3*	187
AT3G15210	*ERF4*	*ERF-B1*	40
AT2G44940		*DREB-A4*	47
AT2G20880	*ERF53*	*ERF*	33
AT5G18450		*DREB-A2*	74
AT1G71450		*DREB-A4*	39
AT5G21960		*DREB-A5*	26
AT1G64380		*DREB-A6*	79
AT1G28360	*ERF12*	*ERF-B1*	96
AT1G53170	*ERF8*	*ERF-B1*	78
AT5G51190		*ERF-B3*	8
AT5G25190	*ESE3*	*ERF-B6*	81
AT5G11190	*SHN2*	*ERF-B6*	37
Total			1970

Interaction network constructions

From the interaction results, we realized that some preys were identified in screens with two or more baits; this is reasonable because the baits are members of the same AP2/EREBP family. To assess the full extent of parallel interactions of AP2/EREBP proteins, we counted the same preys that interacted with different bait proteins and identified 300 pairs of interactions that show the pattern of the same prey interacting with various baits. We organized these bait-prey pairs into five groups on the basis of the proteins functions including 26 proteins

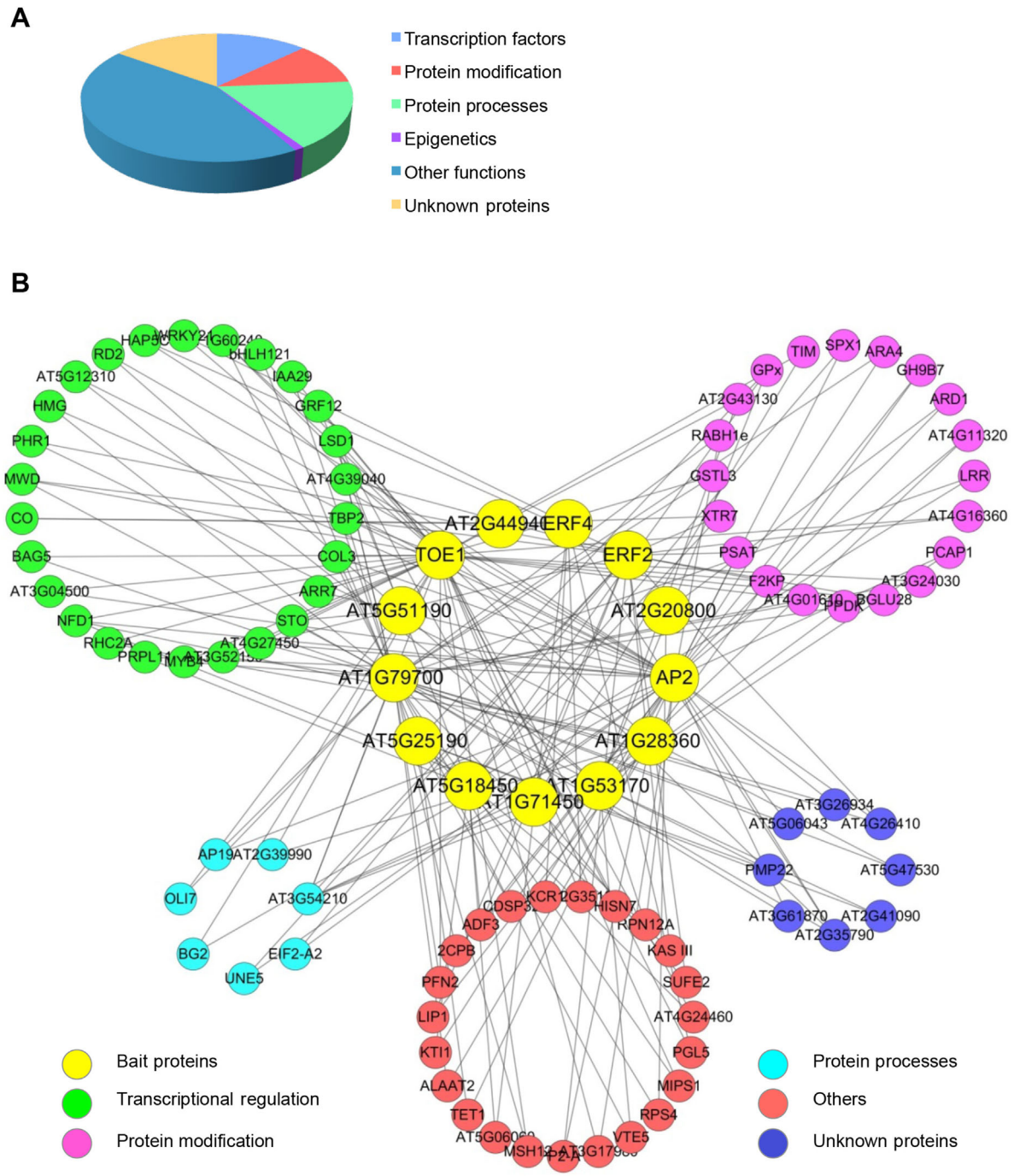

Figure 4. Functional classifications of prey proteins and core interaction network of AP2/EREBP proteins
(A) Functional classification of the prey proteins. **(B)** Core interaction networks constructed using Cytoscape. All of the preys shown here interact with at least two baits. We divided the proteins into five groups including bait proteins, transcriptional regulation, protein modification, protein processes, unknown proteins and others. Proteins are represented by circles, interactions are indicated by lines.

annotated for transcriptional regulation, 20 proteins for protein modification, seven proteins predicted for protein processes, nine unknown proteins and 23 proteins with possibly other functions (Figure 4B).

The AP2/EREBP proteins act (potentially) as transcriptional factors that bind to target DNA and recruit other proteins to regulate transcription. Our results suggest that there are 252 transcription factors that could interact with one or more of the selected 16 baits, implying that AP2/EREBP proteins might form transcriptional complexes with other transcription factors. Other non-transcription factor proteins might participate in the function and/or regulation of the AP2/EREBP

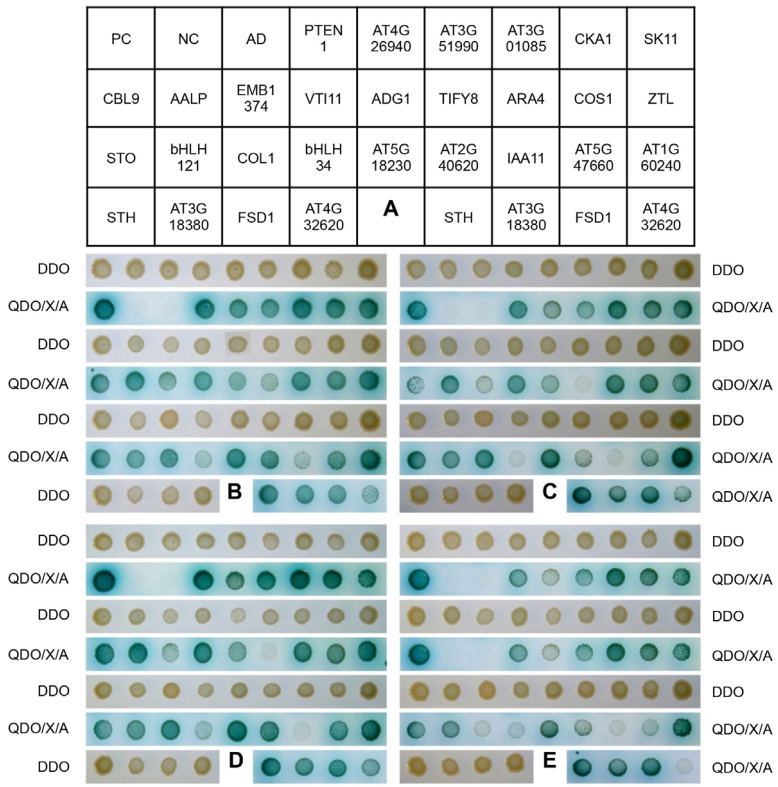

Figure 5. Retests of the interaction of TOE1 preys with TOE1, TOE2, TOE3 and SNZ in yeast cells
(A) Schematic diagram of yeast patches containing the prey proteins of TOE1 that are tested in results shown in **(B)**, **(C)**, **(D)** and **(E)**. Yeast two-hybrid validation of the interactions between TOE1-prey-proteins and TOE1 **(B)**, TOE2 **(C)**, TOE3 **(D)** and SNZ **(E)**. DDO, mediam lacking Trp, Leu; QDO/X/A median lacking His, Ade, Trp, Leu and with Aba and X-α-gal. Blue color indicates interaction.

proteins in another fashion; this will need further studies to understand.

Gene ontology analyses

To gain some additional clues about the biological significance of the interactions between AP2/EREBP proteins and their prey proteins, we analyzed the gene ontology of the prey proteins. We classified the prey proteins according to their putative functions in three ways, including cellular components, biological processes and molecular functions. With regard to cellular components, most of the prey proteins are predicted to localize in the nucleus, strongly suggesting that the prey proteins function in the same cellular structures as their partners (Figure 7A). For biological processes, most of the prey proteins are annotated to participate in other cellular processes and metabolic processes. In addition, the prey proteins also predicted to participate in developmental processes, response to abiotic, biotic stimuli and various stresses (Figure 7B). Also, transcription and signal transduction were enriched in prey proteins. Finally, the molecular function enrichment showed that the prey proteins participate in various binding functions including nucleotide (DNA or RNA) binding, and protein binding, as well as kinase, transcription factor and other activities (Figure 7C). Therefore, several predicted functions of the interactive partners of AP2/EREBPs are functionally consistent with their baits.

Effect of anther developmental regulators on expression of *AP2* family members

The putative protein–protein interactions reported above suggest that members of the *AP2* family with known developmental functions could be linked to regulators of environmental responses, and vice versa. Another way regulators of development and environmental responses could affect each other is through gene expression. As an illustration, we examined transcriptome datasets generated in our lab from wild-type and mutant anthers (Zhu et al. 2015) (Tables 3, S1) for possible effects of the mutations on the expression of *AP2* family members. *DYT1* is a key regulator of anther gene expression and the DYT1 protein interacts with three closely-related bHLH proteins, bHLH010, bHLH089, and bHLH091 (Zhang et al. 2006; Feng et al. 2012). Our recent experiments indicate that a triple mutant of the three bHLH genes is severely defective in anther development, with reduced expression of known anther developmental genes (Zhu et al. 2015). We found that among the *AP2* family members, 37 had reduced expression in the *bhlh* triple mutant anthers, compared with wild-type anthers (Table 3), suggesting that they are regulated positively by these bHLH proteins. Among the 37 *AP2* family genes, five were also affected in the *dyt1* mutant (Table S1), further supporting the idea that members of the *AP2* family are regulated by key anther

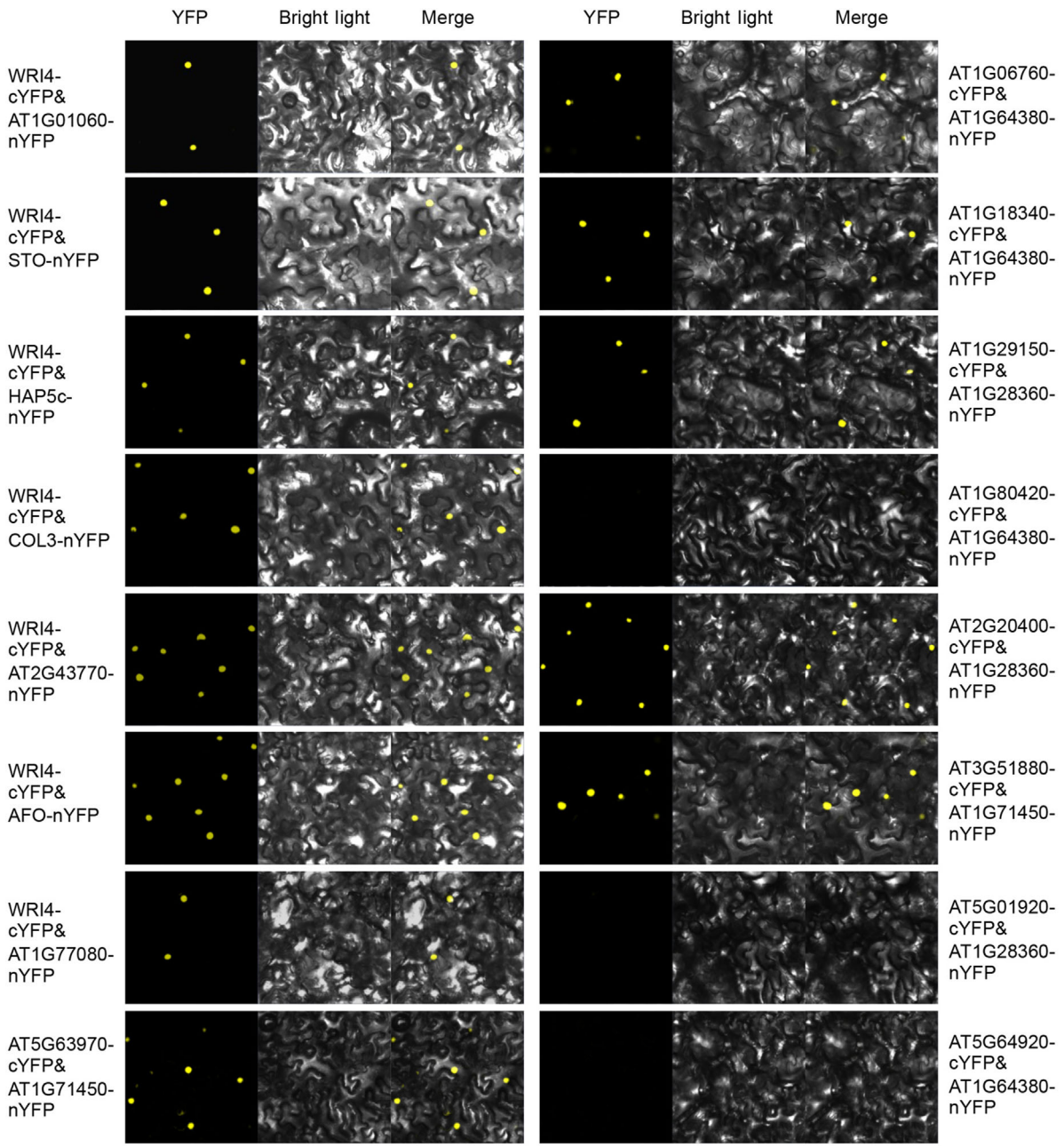

Figure 6. Validation of the interactions *in planta* by BiFC
Bimolecular fluorescence complementation (BiFC) analysis was performed to validate the interaction between selected baits and prey proteins. All baits were fused with cYFP; all preys were fused with nYFP. YFP, fluorescence of yellow fluorescent protein; Bright, bright field; Merge, merge of YFP and Bright. The bright dot(s) in the YFP panels represent interaction.

developmental genes. It is worth noting that many of the affected *AP2* family members are annotated to have functions in environmental responses in the *EREBP* subfamily, including the *CBF1/2/3* genes important for cold tolerance (Stockinger et al. 1997; Gilmour et al. 1998; Medina et al. 1999). Thus, the positive effects on the expression of these known/putative mediators of environmental signal by the three bHLH factors and DYT1 provide an additional link between regulators of development and environmental responses.

DISCUSSION

The AP2/EREBP proteins have important functions in regulating plant development and environmental responses. Our phylogenetic analyses indicate that the *AP2/EREBP* genes in Brassicaceae exhibited similar patterns among subfamilies: (i) the expansion of *AP2/EREBP* genes occurred in the ancestor of Brassicaceae; in *B. rapa*, which has experienced a recent WGD, additional expansion has occurred; (ii) most of the gene losses following the alpha and beta WGDs had

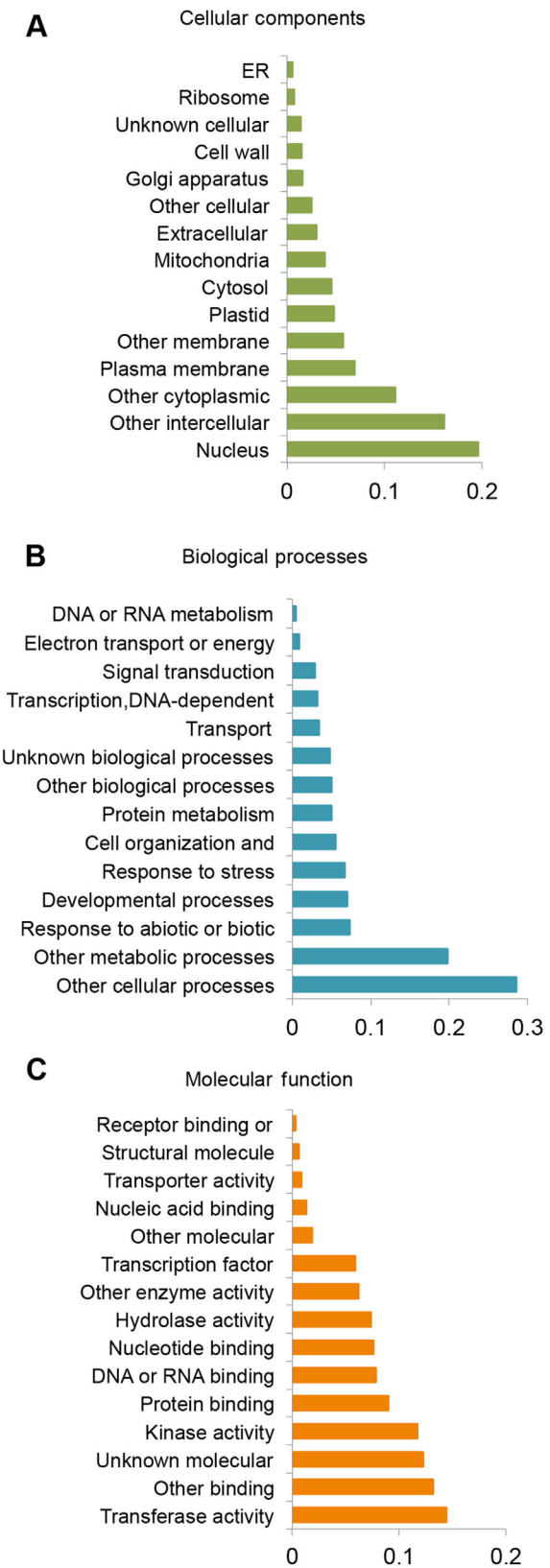

Figure 7. Gene ontology analyses of prey proteins
Fractions of prey proteins in GO categories according to the cellular components **(A)**, biological processes **(B)**; and molecular function **(C)**.

Table 3. *AP2* genes affected in the *bhlh triple* mutant

Gene*	Alias	GFOLD**	RPKM Col-0	RPKM *triple***
AT1G12610	DDF1	−6.25996	49.4785	0.144581
AT1G13260	RAV1	−1.88405	10.4447	0.887313
AT1G19210		−1.95568	6.01208	0.438025
AT1G22190	RAP2.4	−2.85424	47.8332	2.91085
AT1G25560	TEM1	−1.29225	27.781	5.27553
AT1G28360	ERF12	−1.61463	55.6178	7.71617
AT1G28370	ERF11	−3.48987	12.5389	0.379413
AT1G43160	RAP2.6	−6.62988	40.3182	0
AT1G46768	RAP2.1	−1.23229	1.54851	0.116088
AT1G53170	ERF8	−1.51007	137.639	19.7702
AT1G63030	DDF2	−3.6206	4.18268	0
AT1G74930	ORA47	−3.39263	46.6051	2.0298
AT1G75490		−5.06368	13.3635	0.0966472
AT1G79700	WRI4	−1.14185	52.132	9.67008
AT2G20880	ERF53	−5.03813	14.5766	0.105244
AT2G38340	DREB19	−1.29931	5.49735	0.758202
AT2G44840	ERF13	−1.72266	16.2954	1.70774
AT3G15210	ERF4	−1.39304	146.195	25.5851
AT3G61630	CRF6	−1.04539	28.3421	5.94682
AT4G17490	ERF6	−1.06453	6.92583	1.28284
AT4G17500	ERF1	−3.00152	20.7286	1.14626
AT4G25470	CBF2	−4.05797	25.2195	0.534597
AT4G25480	CBF3	−2.03661	39.4682	3.63403
AT4G25490	CBF1	−5.33528	26.5346	0.123108
AT4G28140		−5.12617	25.8455	0.207167
AT4G34410	ERF109	−3.09166	62.0548	3.18127
AT4G36900	DEAR4	−1.20982	17.4523	3.41191
AT5G05410	DREB2A	−3.29537	131.018	6.51857
AT5G13330	RAP2.6L	−1.60165	33.5581	4.56017
AT5G21960		−1.94519	3.13743	0.246215
AT5G47220	ERF2	−4.08149	22.286	0.544771
AT5G51190	ERF105	−1.58481	5.00341	0.629135
AT5G51990	CBF4	−5.02734	30.3583	0.287555
AT5G52020		−2.56442	1.72051	0
AT5G61600	ERF104	−2.21398	35.3816	3.21333
AT5G64750	ABR1	−1.59447	3.87442	0.515635
AT5G67190	DEAR2	−1.33471	13.028	2.02843

*Genes in boldface were the baits in the Y2H screening with positive clones, and those in italic were down-regulated in the *dyt1-3* mutant as well. **GFOLD value could be recognized as the adjusted log2(Fold change), see Feng J, Meyer CA, Wang Q, Liu JS, Liu XS, Zhang Y. GFOLD: a generalized fold change for ranking differentially expressed genes from RNA-seq data. **Bioinformatics** 2012. ***_triple_ is an abbreviation for the *bhlh010 bhlh089 amiR-bHLH091* triple mutant.

occurred before the divergence of the Brassicaceae species examined here; (iii) among the ancestral *AP2/EREBP* genes, some of them have lost all but one copy, while others have retained two or three copies, but few retained all four copies. Our protein–protein interaction results of AP2/EREBP proteins provided an informative dataset for testing regulatory mechanisms and also afforded new insights for uncharacteristic biological pathways. In addition, our results also suggested that the AP2 subfamily proteins also interact

with proteins related to stress response and the EREBP subfamily proteins interact with developmental associated proteins. Therefore, earlier analyses of genes with two AP2 domains (*AP2*-like) involved in plant development and genes with one AP2 domain (*EREBP*s) related to environmental response probably represent a subset of their functions. Rather, the results here suggest that plant developmental regulation is likely coupled with pathways for environmental information.

Previous studies have reported interactomes for specific proteins, including those for G-protein and TOPLESS in *Arabidopsis* and MAPK and MAPKK in rice (Klopffleisch et al. 2011; Causier et al. 2012; Singh et al. 2012; Wankhede et al. 2013), but not yet for multiple members of a transcription factor family. Our results on the potential protein–protein interaction of representative members of the AP2/EREBP family provide useful information for further functional studies.

Several *AP2/EREBP* genes have been well studied in *A. thaliana*, and our selected baits for Y2H screens included many of them. AP2 is known to regulate floral organ identity; two AP2-interacting proteins are key regulators for brassinosteroid signaling, the transcription factors BES1/BZR1 (bri1-Ethylmethane Sulphonate suppressor 1/brassinazole-resistant 1) and the 14-3-3 protein GRF10. BR was found to regulate male fertility by affecting the expression of the anther and pollen development genes (Ye et al. 2010). Thus, our results of AP2-BES1 and AP2-GRF10 interactions might provide information on regulatory mechanisms linking organ identity to other aspects of floral development.

The WRI4 protein was reported to function as a transcription factor to regulate lipid biosynthetic pathways in during *Arabidopsis* development (To et al. 2012). Our Y2H screening with WRI4 identified over 50 clones for the same prey: the B-BOX protein CO-like 3 (COL3); furthermore, this interaction was verified by a BiFC experiment. A previous study showed that COL3 positively regulates seedling growth under light and shoot branching in a day-length sensitive manner (Datta et al. 2006). Our interaction results suggest that WRI4-dependent processes, such as lipid biosynthesis, might be regulated by light or day-length via a COL3-mediated signaling pathway.

The ERF-B1 protein ERF4 and DREB-A4 protein AT2G44940 can both interact with CONSTANS (CO) in yeast. Although *AT2G44940* has not been studied genetically, the *ERF4* gene acts as transcriptional repressor negatively modulating ethylene and abscisic acid responses (Yang et al. 2005). CO is required for long-day dependent flowering by activating the expression of the *FT* gene encoding a flowering signal (Amasino 2010; Song et al. 2012). However, previous studies have not revealed an interaction between these two type proteins. Flowering time control is modulated by the interactions of endogenous developmental processes and exogenous environmental signals (Amasino 2010; Song et al. 2010). As a central regulator of photoperiodic flowering, the abundance of the CO protein and the regulation of its activity are important for flowering time regulation. Consistently, CO could interact with dozens of proteins in modulating its protein stability and activity (Valverde 2011), as further supported by our recent findings that TOE1 interacts with CO and inhibits its activity (Zhang et al. 2015). In addition, it is known that environmental stresses, such as water availability, can trigger early flowering. Our finding suggests that stress-responsive regulators ERF4 and AT2G44940 might convey the stress signals to the CO protein to influence photoperiodic flowering.

MATERIALS AND METHODS

Data collection of *AP2/EREBP* family genes

Whole-genome annotated sequences of *Arabidopsis thaliana*, *Arabidopsis lyrata*, *Capsella rubella*, *Brassica rapa*, *Eutrema salsugineum* and *Populus trichocarpa* were downloaded from public databases of phytozome (http://www.phytozome.net/). Then the HMMER (profile hidden Markov models) tool (v3.1) (Finn et al. 2011) was carried out using the entry *AP2* (PF00847) of Pfam-A downloaded from Pfam (v27.0) (http://pfam.sanger.ac.uk/) to identify *AP2* homologous genes from each of the above datasets. To verify the obtained sequences were *AP2* homologs, all retrieved proteins were searched for the AP2 domain using the Pfam tool (http://pfam.sanger.ac.uk/) and SMART (http://smart.embl-heidelberg.de/) with E values setting below 1e-5, and sequences lacking of the AP2 domain were discarded for further analyses.

Phylogenetic analyses of *AP2/EREBP* gene family

Multiple amino acid sequences were firstly aligned by MUSCLE (v3.8.31) (Edgar 2004) using default settings. For improving the alignment of variable regions in 5' and 3' ends of genes, a preliminary phylogenetic tree was constructed to divide sequences into several subgroups. Sequences in each subgroup were further aligned by MUSCLE (v3.8.31) and then manually adjusted using MEGA (v5). All aligned subgroups were combined into a global alignment by MUSCLE (v3.8.31) with the "profile" setting.

Maximum likelihood (ML) trees of aligned amino acids matrix were constructed using RAxML (v7.0.4) (Stamatakis 2006) with the JTT + CAT model and 100 bootstrap replicates.

Intragenome syntenic relationships of AP2/EREBP family genes of *A. thaliana* were implemented by PGDD (http://chibba.agtec.uga.edu/duplication/) to detect the contribution of genome duplication in the expansion of *AP2/EREBP* family.

Exon numbers and domain conservation analyses

Nucleotide coordinates for the start and end of each exon were extracted from all of the above seven annotated genomes. Exon numbers of each *AP2/EREBP* homolog were then calculated using custom Perl script.

Amino acid sequence of AP2 domain in each AP2/EREBP homolog was fetched by custom Perl script according to parsed results for domains by HMMER. WebLogo (http://weblogo.berkeley.edu/) was applied to generate subfamily specific domain logo for detecting conserved and/or diversified amino acids.

Yeast two-hybrid screening and assays

Using the phylogenetic tree of all the *AP2/EREBP* family members in five Brassicaceae species, 56 *A. thaliana AP2/EREBP* genes either with known functions and/or representing

clades containing members from other Brassicaceae species were chosen as baits. These bait genes cover one-third of the AP2/EREBP family in *A. thaliana*, and distribute in almost all sub-groups for a general representation. The Y2H screening was performed according to Matchmaker Gold Yeast Two-Hybrid System User Manual (Clontech). cDNAs for 50 of the 56 bait genes were successfully cloned into the bait vector pGBKT7 (Clontech), which were introduced Y2H-Gold yeast cells by transformation (Clontech). After the elimination due to autoactivation and toxicity in yeast, 21 bait constructs were used for Y2H screening by mating with *A. thaliana* Y2H cDNA library (Clontech, approximately 1.0×10^7 transformants). The mating mixtures were screened with the growth selection on media lacking His, Ade, Trp, Leu and with Aba and X-α-gal (QDO/X/A). The positive colonies were transferred to new QDO/X/A plates and performed yeast colony PCR to obtain the sequences of the prey. For Y2H assays, the full-length cDNAs were cloned into pGBKT7 or pGADT7 and introduced into the Y2H-Gold or Y187 yeast strains, respectively. Then the two yeast strains were mated and transferred to the DDO and QDO/X/A plate for selection. Information for all primers in this study is provided in the supplemental materials (Table S2).

Bimolecular fluorescence complementation assays

cDNAs of bait genes were cloned into the pXY104 vector. cDNAs of prey genes were cloned into the pXY106 vector. The resulting cassettes including fusion proteins and constitutive promoters were transformed into *Agrobacterium*. For BiFC experiments, 3-week-old tobacco (*Nicotiana benthamiana*) leaves were co-infiltrated with *Agrobacterium* containing relevant constructs. After 36–48 h, expressions of various fluorescent proteins were analyzed by confocal microscopy (Zeiss).

Interaction network construction

To illustrate interactions between baits and preys detected by Y2H screening and partially confirmed with BiFC, the interaction network constructing software Cytoscape was applied to construct primary interaction networks of the *A. thaliana* AP2/EREBP members in this study.

ACKNOWLEDGEMENTS

The authors thank Professor Jianxiang Liu for providing the BiFC plasmids and Engao Zhu and Fang Chang for sharing results before publication, and gratefully acknowledge financial support from the National Natural Science Foundation of China (91131007) and funds from Fudan University.

AUTHOR CONTRIBUTIONS

H.M. and B.Z. designed the study. L.Z. carried out evolutionary analysis. L.Z. and C.Y. carried out the bioinformatics analysis. B.Z. and Y.Y. performed the Y2H screening experiment. B.Z. performed the Y2H confirmation experiment. B.Z., Y.Y., Q.P., T.J. and D.X. performed BiFC experiments. B.Z., L.Z. and Y.Y. analyzed the results. B.Z. and L.Z. drafted the manuscript and H.M. revised the manuscript. All authors read and approved the final manuscript.

REFERENCES

Amasino R (2010) Seasonal and developmental timing of flowering. **Plant J** 61: 1001–1013

Arabidopsis Interactome Mapping C (2011) Evidence for network evolution in an *Arabidopsis* interactome map. **Science** 333: 601–607

Aukerman MJ, Sakai H (2003) Regulation of flowering time and floral organ identity by a microRNA and its *APETALA2*-like target genes. **Plant Cell** 15: 2730–2741

Blanc G, Hokamp K, Wolfe KH (2003) A recent polyploidy superimposed on older large-scale duplications in the *Arabidopsis* genome. **Genome Res** 13: 137–144

Bowman JL, Smyth DR, Meyerowitz EM (1989) Genes directing flower development in *Arabidopsis*. **Plant Cell** 1: 37–52

Bowman JL, Smyth DR, Meyerowitz EM (1991) Genetic interactions among floral homeotic genes of *Arabidopsis*. **Development** 112: 1–20

Causier B, Ashworth M, Guo W, Davies B (2012) The TOPLESS interactome: A framework for gene repression in *Arabidopsis*. **Plant Physiol** 158: 423–438

Chen X (2004) A microRNA as a translational repressor of *APETALA2* in *Arabidopsis* flower development. **Science** 303: 2022–2025

Coen ES, Meyerowitz EM (1991) The war of the whorls: Genetic interactions controlling flower development. **Nature** 353: 31–37

Dassanayake M, Oh DH, Haas JS, Hernandez A, Hong H, Ali S, Yun DJ, Bressan RA, Zhu JK, Bohnert HJ, Cheeseman JM (2011) The genome of the extremophile crucifer *Thellungiella parvula*. **Nat Genet** 43: 913–918

Datta S, Hettiarachchi GH, Deng XW, Holm M (2006) *Arabidopsis CONSTANS-LIKE3* is a positive regulator of red light signaling and root growth. **Plant Cell** 18: 70–84

Dietz KJ, Vogel MO, Viehhauser A (2010) AP2/EREBP transcription factors are part of gene regulatory networks and integrate metabolic, hormonal and environmental signals in stress acclimation and retrograde signalling. **Protoplasma** 245: 3–14

Edgar RC (2004) MUSCLE: Multiple sequence alignment with high accuracy and high throughput. **Nucleic Acids Res** 32: 1792–1797

Elliott RC, Betzner AS, Huttner E, Oakes MP, Tucker WQ, Gerentes D, Perez P, Smyth DR (1996) *AINTEGUMENTA*, an *APETALA2*-like gene of Arabidopsis with pleiotropic roles in ovule development and floral organ growth. **Plant Cell** 8: 155–168

Feng B, Lu D, Ma X, Peng Y, Sun Y, Ning G, Ma H (2012) Regulation of the *Arabidopsis* anther transcriptome by *DYT1* for pollen development. **Plant J** 72: 612–624

Finn RD, Clements J, Eddy SR (2011) HMMER web server: Interactive sequence similarity searching. **Nucleic Acids Res** 39: W29–W37

Fujimoto SY, Ohta M, Usui A, Shinshi H, Ohme-Takagi M (2000) *Arabidopsis* ethylene-responsive element binding factors act as transcriptional activators or repressors of GCC box-mediated gene expression. **Plant Cell** 12: 393–404

Gilmour SJ, Zarka DG, Stockinger EJ, Salazar MP, Houghton JM, Thomashow MF (1998) Low temperature regulation of the *Arabidopsis* CBF family of AP2 transcriptional activators as an early step in cold-induced *COR* gene expression. **Plant J** 16: 433–442

Hu TT, Pattyn P, Bakker EG, Cao J, Cheng JF, Clark RM, Fahlgren N, Fawcett JA, Grimwood J, Gundlach H, Haberer G, Hollister JD, Ossowski S, Ottilar RP, Salamov AA, Schneeberger K, Spannagl M, Wang X, Yang L, Nasrallah ME, Bergelson J, Carrington JC, Gaut BS, Schmutz J, Mayer KF, Van de Peer Y, Grigoriev IV, Nordborg M, Weigel D, Guo YL (2011) The *Arabidopsis lyrata* genome sequence

and the basis of rapid genome size change. **Nat Genet** 43: 476–481

Hu YX, Wang YX, Liu XF, Li JY (2004) *Arabidopsis RAV1* is down-regulated by brassinosteroid and may act as a negative regulator during plant development. **Cell Res** 14: 8–15

Iwase A, Mitsuda N, Koyama T, Hiratsu K, Kojima M, Arai T, Inoue Y, Seki M, Sakakibara H, Sugimoto K, Ohme-Takagi M (2011) The *AP2/ERF* transcription factor *WIND1* controls cell dedifferentiation in *Arabidopsis*. **Curr Biol** 21: 508–514

Jiao Y, Wickett NJ, Ayyampalayam S, Chanderbali AS, Landherr L, Ralph PE, Tomsho LP, Hu Y, Liang H, Soltis PS, Soltis DE, Clifton SW, Schlarbaum SE, Schuster SC, Ma H, Leebens-Mack J, dePamphilis CW (2011) Ancestral polyploidy in seed plants and angiosperms. **Nature** 473: 97–100

Jofuku KD, den Boer BG, Van Montagu M, Okamuro JK (1994) Control of *Arabidopsis* flower and seed development by the homeotic gene *APETALA2*. **Plant Cell** 6: 1211–1225

Jofuku KD, Omidyar PK, Gee Z, Okamuro JK (2005) Control of seed mass and seed yield by the floral homeotic gene *APETALA2*. **Proc Natl Acad Sci USA** 102: 3117–3122

Kim S, Soltis PS, Wall K, Soltis DE (2006) Phylogeny and domain evolution in the *APETALA2*-like gene family. **Mol Biol Evol** 23: 107–120

Kizis D, Lumbreras V, Pages M (2001) Role of AP2/EREBP transcription factors in gene regulation during abiotic stress. **FEBS Lett** 498: 187–189

Klopffleisch K, Phan N, Augustin K, Bayne RS, Booker KS, Botella JR, Carpita NC, Carr T, Chen JG, Cooke TR, Frick-Cheng A, Friedman EJ, Fulk B, Hahn MG, Jiang K, Jorda L, Kruppe L, Liu C, Lorek J, McCann MC, Molina A, Moriyama EN, Mukhtar MS, Mudgil Y, Pattathil S, Schwarz J, Seta S, Tan M, Temp U, Trusov Y, Urano D, Welter B, Yang J, Panstruga R, Uhrig JF, Jones AM (2011) *Arabidopsis* G-protein interactome reveals connections to cell wall carbohydrates and morphogenesis. **Mol Syst Biol** 7: 532

Klucher KM, Chow H, Reiser L, Fischer RL (1996) The *AINTEGUMENTA* gene of *Arabidopsis* required for ovule and female gametophyte development is related to the floral homeotic gene *APETALA2*. **Plant Cell** 8: 137–153

Krizek BA (2003) *AINTEGUMENTA* utilizes a mode of DNA recognition distinct from that used by proteins containing a single AP2 domain. **Nucleic Acids Res** 31: 1859–1868

Kunst L, Klenz JE, Martinez-Zapater J, Haughn GW (1989) *AP2* gene determines the identity of perianth organs in flowers of *Arabidopsis thaliana*. **Plant Cell** 1: 1195–1208

Mathieu J, Yant LJ, Mürdter F, Küttner F, Schmid M (2009) Repression of flowering by the miR172 target *SMZ*. **PLoS Biol** 7: e1000148

Medina J, Bargues M, Terol J, Perez-Alonso M, Salinas J (1999) The *Arabidopsis CBF* gene family is composed of three genes encoding AP2 domain-containing proteins whose expression is regulated by low temperature but not by abscisic acid or dehydration. **Plant Physiol** 119: 463–470

Mizoi J, Shinozaki K, Yamaguchi-Shinozaki K (2012) AP2/ERF family transcription factors in plant abiotic stress responses. **Biochim Biophys Acta** 1819: 86–96

Ohto MA, Fischer RL, Goldberg RB, Nakamura K, Harada JJ (2005) Control of seed mass by *APETALA2*. **Proc Natl Acad Sci USA** 102: 3123–3128

Okamuro JK, Caster B, Villarroel R, Van Montagu M, Jofuku KD (1997) The AP2 domain of APETALA2 defines a large new family of DNA binding proteins in *Arabidopsis*. **Proc Natl Acad Sci USA** 94: 7076–7081

Riechmann JL, Meyerowitz EM (1998) The AP2/EREBP family of plant transcription factors. **Biol Chem** 379: 633–646

Sakuma Y, Liu Q, Dubouzet JG, Abe H, Shinozaki K, Yamaguchi-Shinozaki K (2002) DNA-binding specificity of the ERF/AP2 domain of *Arabidopsis* DREBs, transcription factors involved in dehydration- and cold-inducible gene expression. **Biochem Biophys Res Commun** 290: 998–1009

Schmid M, Uhlenhaut NH, Godard F, Demar M, Bressan R, Weigel D, Lohmann JU (2003) Dissection of floral induction pathways using global expression analysis. **Development** 130: 6001–6012

Shigyo M, Ito M (2004) Analysis of gymnosperm two-AP2-domain-containing genes. **Dev Genes Evol** 214: 105–114

Singh R, Lee MO, Lee JE, Choi J, Park JH, Kim EH, Yoo RH, Cho JI, Jeon JS, Rakwal R, Agrawal GK, Moon JS, Jwa NS (2012) Rice mitogen-activated protein kinase interactome analysis using the yeast two-hybrid system. **Plant Physiol** 160: 477–487

Slotte T, Hazzouri KM, Agren JA, Koenig D, Maumus F, Guo YL, Steige K, Platts AE, Escobar JS, Newman LK, Wang W, Mandakova T, Vello E, Smith LM, Henz SR, Steffen J, Takuno S, Brandvain Y, Coop G, Andolfatto P, Hu TT, Blanchette M, Clark RM, Quesneville H, Nordborg M, Gaut BS, Lysak MA, Jenkins J, Grimwood J, Chapman J, Prochnik S, Shu S, Rokhsar D, Schmutz J, Weigel D, Wright SI (2013) The *Capsella rubella* genome and the genomic consequences of rapid mating system evolution. **Nat Genet** 45: 831–835

Somers DE, Kim WY, Geng R (2004) The F-box protein ZEITLUPE confers dosage-dependent control on the circadian clock, photomorphogenesis, and flowering time. **Plant Cell** 16: 769

Somers DE, Schultz TF, Milnamow M, Kay SA (2000) *ZEITLUPE* encodes a novel clock-associated PAS protein from *Arabidopsis*. **Cell** 101: 319–329

Song YH, Ito S, Imaizumi T (2010) Similarities in the circadian clock and photoperiodism in plants. **Curr Opin Plant Biol** 13: 594–603

Song YH, Smith RW, To BJ, Millar AJ, Imaizumi T (2012) FKF1 conveys timing information for CONSTANS stabilization in photoperiodic flowering. **Science** 336: 1045–1049

Stamatakis A (2006) RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. **Bioinformatics** 22: 2688–2690

Stockinger EJ, Gilmour SJ, Thomashow MF (1997) *Arabidopsis thaliana CBF1* encodes an AP2 domain-containing transcriptional activator that binds to the C-repeat/DRE, a cis-acting DNA regulatory element that stimulates transcription in response to low temperature and water deficit. **Proc Natl Acad Sci USA** 94: 1035–1040

Takase T, Nishiyama Y, Tanihigashi H, Ogura Y, Miyazaki Y, Yamada Y, Kiyosue T (2011) *LOV KELCH PROTEIN2* and *ZEITLUPE* repress *Arabidopsis* photoperiodic flowering under non-inductive conditions, dependent on *FLAVIN-BINDING KELCH REPEAT F-BOX1*. **Plant J** 67: 608–621

To A, Joubes J, Barthole G, Lecureuil A, Scagnelli A, Jasinski S, Lepiniec L, Baud S (2012) *WRINKLED* transcription factors orchestrate tissue-specific regulation of fatty acid biosynthesis in *Arabidopsis*. **Plant Cell** 24: 5007–5023

Valverde F (2011) *CONSTANS* and the evolutionary origin of photoperiodic timing of flowering. **J Exp Bot** 62: 2453–2463

Wang X, Wang H, Wang J, Sun R, Wu J, Liu S, Bai Y, Mun JH, Bancroft I, Cheng F, Huang S, Li X, Hua W, Wang J, Wang X, Freeling M, Pires JC, Paterson AH, Chalhoub B, Wang B, Hayward A, Sharpe AG, Park BS, Weisshaar B, Liu B, Li B, Liu B, Tong C, Song C, Duran C, Peng C, Geng C, Koh C, Lin C, Edwards D, Mu D, Shen D, Soumpourou E, Li F, Fraser F, Conant G, Lassalle G, King GJ, Bonnema G, Tang H, Wang H, Belcram H, Zhou H, Hirakawa H, Abe

H, Guo H, Wang H, Jin H, Parkin IA, Batley J, Kim JS, Just J, Li J, Xu J, Deng J, Kim JA, Li J, Yu J, Meng J, Wang J, Min J, Poulain J, Wang J, Hatakeyama K, Wu K, Wang L, Fang L, Trick M, Links MG, Zhao M, Jin M, Ramchiary N, Drou N, Berkman PJ, Cai Q, Huang Q, Li R, Tabata S, Cheng S, Zhang S, Zhang S, Huang S, Sato S, Sun S, Kwon SJ, Choi SR, Lee TH, Fan W, Zhao X, Tan X, Xu X, Wang Y, Qiu Y, Yin Y, Li Y, Du Y, Liao Y, Lim Y, Narusaka Y, Wang Y, Wang Z, Li Z, Wang Z, Xiong Z, Zhang Z, Brassica rapa Genome Sequencing Project C (2011) The genome of the mesopolyploid crop species *Brassica rapa*. **Nat Genet** 43: 1035–1039

Wankhede DP, Misra M, Singh P, Sinha AK (2013) Rice mitogen activated protein kinase kinase and mitogen activated protein kinase interaction network revealed by in-silico docking and yeast two-hybrid approaches. **PLoS ONE** 8: e65011

Weigel D (1995) The APETALA2 domain is related to a novel type of DNA binding domain. **Plant Cell** 7: 388–389

Yan H, Marquardt K, Indorf M, Jutt D, Kircher S, Neuhaus G, Rodriguez-Franco M (2011) Nuclear localization and interaction with COP1 are required for STO/BBX24 function during photomorphogenesis. **Plant Physiol** 156: 1772–1782

Yang Z, Tian L, Latoszek-Green M, Brown D, Wu K (2005) *Arabidopsis* ERF4 is a transcriptional repressor capable of modulating ethylene and abscisic acid responses. **Plant Mol Biol** 58: 585–596

Yant L, Mathieu J, Dinh TT, Ott F, Lanz C, Wollmann H, Chen X, Schmid M (2010) Orchestration of the floral transition and floral development in *Arabidopsis* by the bifunctional transcription factor APETALA2. **Plant Cell** 22: 2156–2170

Ye Q, Zhu W, Li L, Zhang S, Yin Y, Ma H, Wang X (2010) Brassinosteroids control male fertility by regulating the expression of key genes involved in *Arabidopsis* anther and pollen development. **Proc Natl Acad Sci USA** 107: 6100–6105

Zhang B, Wang L, Zeng L, Zhang C, Ma H (2015) Arabidopsis TOE proteins convey a photoperiodic signal to antagonize CONSTANS and regulate flowering time. **Genes Dev** 29: 975–987

Zhang W, Sun Y, Timofejeva L, Chen C, G rossniklaus U, Ma H(2006) Regulation of *Arabidopsis* tapetum development and function by *DYSFUNCTIONAL TAPETUM1 (DYT1)* encoding a putative bHLH transcription factor. **Development** 133: 3085–3095

Zhu E, You C, Wang S, Cui J, Niu B, Wang Y, Qi J, Ma H, Fang Chang F (2015) The DYT1-interacting proteins bHLH010, bHLH089 and bHLH091 are redundantly required for *Arabidopsis* anther development and transcriptome. **Plant J** 83: 976–990

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's web-site.

Figure S1. A phylogenetic tree of *DREB-A2* subfamily with several *DREB-A6* genes as outgroup
Bootstrap supports (>50) obtained by RAxML are shown. Gene IDs of each species are shown in different colors. Black circles represent the common ancestor of rosids; blue stars stand for tandem duplications in *Arabidopsis thaliana*; red triangles indicate retained genes after two rounds of WGDs in the ancestor of Brassicaceae.

Figure S2. A phylogenetic tree of *DREB-A4* subfamily with several *DREB-A1* genes as outgroup
Bootstrap supports (>50) obtained by RAxML are shown. Gray circles represent that the common ancestor of rosids contain fourteen ancestor *DREB-A4* genes; black triangles indicate retained genes after two rounds of WGDs in early Brassicaceae.

Figure S3. A phylogenetic tree of *DREB-A5* subfamily with several *DREB-A4* genes as outgroup
Bootstrap supports (>50) obtained by RAxML are shown
Gray circles represent that the common ancestor of rosids contain ten ancestor *DREB-A5* genes; black triangles indicate retained genes after two rounds of WGDs in early Brassicaceae.

Figure S4. A phylogenetic tree of *DREB-A6* subfamily with several *DREB-A2* genes as outgroup
Bootstrap supports (>50) obtained by RAxML are shown. Gray circles represent that the common ancestor of rosids at least contain five ancestor *DREB-A6* genes; black triangles indicate retained genes after two rounds of WGDs in early Brassicaceae.

Figure S5. A phylogenetic tree of *ERF-B1* subfamily with several *ERF-B6* genes as outgroup
Bootstrap supports (>50) obtained by RAxML are shown. Gray circles represent that the common ancestor of rosids at least contain fourteen ancestor *ERF-B6* genes; black triangles indicate retained genes after two rounds of WGDs in early Brassicaceae.

Figure S6. A phylogenetic tree of *ERF-B2* subfamily with several *ERF-B6* genes as outgroup
Bootstrap supports (>50) obtained by RAxML are shown. Black circles represent that the common ancestor of rosids contain six ancestor *ERF-B6* genes; red triangles indicate retained genes after two rounds of WGDs in early Brassicaceae.

Figure S7. A phylogenetic tree of *ERF-B3* subfamily with several *ERF-B6* genes as outgroup
Bootstrap supports (>50) obtained by RAxML are shown. Gray circles represent that the common ancestor of rosids at least contain six ancestor *ERF-B6* genes; black triangles indicate retained genes after two rounds of WGDs in early Brassicaceae.

Figure S8. A phylogenetic tree of *ERF-B4* subfamily with several *ERF-B6* genes as outgroup
Bootstrap supports (>50) obtained by RAxML are shown. Gray circles represent that the common ancestor of rosids contain six ancestor *ERF-B4* genes; black triangles indicate retained genes after two rounds of WGDs in early Brassicaceae.

Figure S9. A phylogenetic tree of *ERF-B5* subfamily with several *ERF-B6* genes as outgroup
Bootstrap supports (>50) obtained by RAxML are shown. Gray circles represent that the common ancestor of rosids at least contain nine ancestor *ERF-B5* genes; black triangles indicate retained genes after two rounds of WGDs in early Brassicaceae.

Figure S10. A phylogenetic tree of *ERF-B6* subfamily with several *ERF-B5* genes as outgroup
Bootstrap supports (>50) obtained by RAxML are shown. Gray circles represent that the common ancestor of rosids at least contain thirteen ancestor *ERF-B6* genes; black triangles indicate retained genes after two rounds of WGDs in early Brassicaceae.

Figure S11. A phylogenetic tree of *RAV* subfamily with several *ERF* genes as outgroup
Bootstrap supports (>50) obtained by RAxML are shown. Gray circles represent that the common ancestor of rosids contain five ancestor *RAV* genes; black triangles indicate

retained genes after two rounds of WGDs in early Brassicaceae.

Figure S12. A phylogenetic tree of *AP2* subfamily with several *ERF* genes as outgroup

Bootstrap supports (>50) obtained by RAxML are shown. Gray circles represent that the common ancestor of rosids contain ancestor *AP2* genes; black triangles figure out retented genes after two rounds of WGDs in the ancestor of Brassicaceae.

Figure S13. Exon numbers of each *AP2*/*EREBP* subfamily genes in *Arabidopsis lyrata* and *Carica papay*

All genes in *AP2* subfamily contained multiple exons (≥6) while most genes in *RAV*, *DREB* and *ERF* subfamilies have single exon or limited exons in both species.

Figure S14. Auto-activation test of the AP2/EREBP proteins in yeast cells

Blue spots stand for severe autoacivation in yeast cells. Left part are Y2H results; right part are bait genes consistent with the left spots.

Figure S15. Auto-activation test of the TOE1 target proteins in yeast cells

(**A**) Schematic diagram of yeast colonies containing the prey proteins of TOE1 are shown in (**B**). (**B**) Yeast two-hybrid validation the interactions between TOE1-pery-proteins and pGBKT7 empty vector.

Table S1. Expression data for *AP2* family members in wild type and mutant anthers

Table S2. Primers sequences used in this study